This workshop introduces the development of model estimation data sets, the structures of the various model components, and the procedures for estimating models. At the end of the series, participants will be able to better manage model development done for them by others (e.g., consultants) and to understand and evaluate the results.

Questions and Answers during Session II

Q: Was there a site we can print this presentation?
A: Yes. Please see Gary Thomas's email titled "Downloads". You can also find the presentation at: http://tmip.fhwa.dot.gov/discussions/webinars/archive/tmw/downloads/presentation_2.pdf

Q: How do you interpret these results and why do you use the percentage sign?
A: Thanks for the question. The probabilities are often expressed as a percentage. This is because, given the levels of service variables, the probabilities indicate the percentage of time that the individual will use that particular mode.
A: So, one way to interpret the probabilities is as follows. Let's say the individual makes this same trip 1000 times. In about 440 of those cases, he will choose to drive alone.

Q: Question about last week's homework, this is for a sequential multinomial logit model, correct?
A: Can you please elaborate what you mean by sequential?
Q: As opposed to a fully nested model...
A: Good point. Yes, this is a simple multinomial logit model. The property of an MNL is that we treat each mode "independent" of the other. In a nested model, we treat some modes "more like each other than the rest". So, for instance, the transit modes are all grouped in a nest.

Q: Basic question: definition of "in-vehicle time" and out of vehicle time"?
A Thanks for bringing this up. In-vehicle time constitutes the time spent in the vehicle or mode of choice. So, for example, if a person walks from home to the bus stop, boards the bus, alights at the destination stop, then walks to her final destination, the in-vehicle time is the time spent on the bus; the out-of-vehicle time is the walk time, wait time etc.
A: We typically differentiate between in-vehicle and out-of-vehicle because people have different sensitivities to in-vehicle and out-of-vehicle times.
A: Usually, the out-of-vehicle components are 2-3 times more onerous than IVTT, as per popular research.

Q: what is PUMS data?
A: PUMS = Public use micro-sample data. You can learn more about these data at: http://www.census.gov/acs/www/Products/PUMS/

Q: What is a "skim matrix"?
A: "Skims" is a term used for the level of service matrices that we generate from highway networks. So, when we say highway skim matrices, we are talking about auto travel times and distances between each zone pair in the model area.
A: Similarly, there could be transit skims such as, in-vehicle times, wait times, transfer times, etc.

Q: Are there any special considerations for dealing with non-traditional households, e.g. same-sex couples?
A: To our knowledge, there haven't been too many such studies. Nonetheless, with the recent advanced travel models, the travel patterns of these households can be captured in a better

fashion. However, there would have to be sufficient data on such households in the survey data set to provide enough information for estimation.
A: Some other residential settlements such as group quarters, dormitories etc. can be treated separately, usually as special generators. This is especially true in large university towns like Madison.
Q: Referring to the previous question, how does it matter? We as modelers are interested to know the number of trips from a HH.
A: You are absolutely right. The question was more about travel patterns of specific types of households. When we use broad cross-class categories, such differences can sometimes be glossed over. This is why we said that in activity models, these households can be accounted for because we are now analyzing each person and each household. The travel patterns will automatically manifest themselves depending on the type of the household.

Q: NHTS?
A: National Household Travel Survey
A: You can find more about NHTS at: http://nhts.ornl.gov/

Q: How do you weight for those without land lines?
A: Very good question. Survey firms are targeting cell-phone users too. So this is one way to include them.

Q: How often is this done?
A: The previous NHTS was conducted in 2001...in addition to a national data set, several urban regions and DOTs purchased add-ons...the previous NHTS before this one was in 1995

Q: What is the difference between deduced and imputed?
A: Let's say a household has three people, and we know that there are two adults and say the # of children variable has a missing value, then you can deduce #children to be 1. But let's say we don't know income. There is no way to deduce it. We can only impute it by looking at similar households.

Q: Weights?
A: As you are aware, HH surveys include only a sample of households. But as a planner or modeler, you want to know the travel patterns of the entire population. So, to bridge the gap between sample and population, we have what are called "expansion weights". Deriving these factors is actually a subject in itself
Q: Refer to Iterative Proportional Fitting, while you develop weights. Consult Wharton School of Business website to know more.
A: Thanks. IPF is definitely one of the more common methods used while developing the weights.

Q: Recreational trips fall into round trips when the person goes for a jog or bike ride!
A: Excellent point. This is a very good example where we need to look at the reported arrival and departure times to deduce what the person actually did. If we just look at the start and end points, we will be forced to think that the person never left home!
Q: For recreation trips such as walking the dog or going for a run/bike ride, wouldn't the O & D be the same? So for a logic check you would want to make sure if O=D that it is a loop trip or trip purpose was a recreation trip.
A: Yes. See response above. In such cases, we will need to look at the start and end times of the trip and also the activity purpose.

Q: In how many years, should surveys be done?
A: This is entirely dependent on the dynamism of the area under study, the budget available to do the surveys etc. Typically, surveys are only done once every 5-10 years.

Q: I didn't get the choice variable discussion would you please reiterate?
A: We will discuss this in greater detail in Session 4. But the "choice" variable refers to the variable we are trying to model. So let's say we are modeling mode choice, and let's say we have four modes: Auto drive alone, Shared ride, transit, non-motorized. Then, the choice variable is Mode = 1,2,3,4 depending on what the chosen mode was.

Q: Will the chat discussion be available later on-line?
A: Yes, we will be posting it.

Q: What was is about the maximum values under creating necessary variables on slide 25?
A: For example, we may want to model auto ownership with choices of 0, 1, 2, 3, and 4+ autos. A new variable would be created with a value equal to the number of autos if it is 3 or less and equal to 4 if the number of autos is 4 or greater.

Q: Is there a good way to estimate future income level data?
A: If we could do that really well, we could make a lot of money in another profession! The best methods would consider trends of growth in real income and knowledge of future development patterns. Remember, we are using incomes in dollars of the base year (e.g. 2005 dollars).

Q: Trip Attraction Question Related to Recreation Trips - In the Case of recreation trips how do you attract them to zones that do not have employment or households (example a state park) Is their a way to do this beside using a special generator.
A: It would have to be a special generator unless you have other variables you could use. For a state park, I can't think of any such variables.

Q: Is Journey to work defined the same as home based work?
A: For this exercise, yes